Project #1: Exploring Weather Trends

Jacob Lowell

4/28/2020

Contents

The objective of this analysis is to compare activity over time between the average temperatures for my nearest large city, Boston, Massachusetts, USA against the average yearly world-wide temperature. To complete this project, I pulled data from the Udacity site with a SQL query, whose results I exported to a .csv file which I then pulled into R for the following analysis which I am able to convert to PDF via R Markdown.

1. SQL

My first step was to write a SQL query to pull the data I needed from the interface on the Udacity site. I live in Freeport, Maine, USA, so the city nearest to me in the data set is Boston, Massachusetts, USA. The SQL query that I used is as follows:

```
select a. * ,
  b.avg_temp as avg_temp_world
  from city_data a  join
  global_data b
  on a.year = b.year
where  a.country = 'United States' and a.city = 'Boston'
```

2 R

For this analysis in R Markdown, I used the tidyverse package for data manipulation and visualization, the tidyquant package for styling and the calculation of moving averages and the GGally package for the correlation plot near the end of the document.

View Boston and World temperature time-series

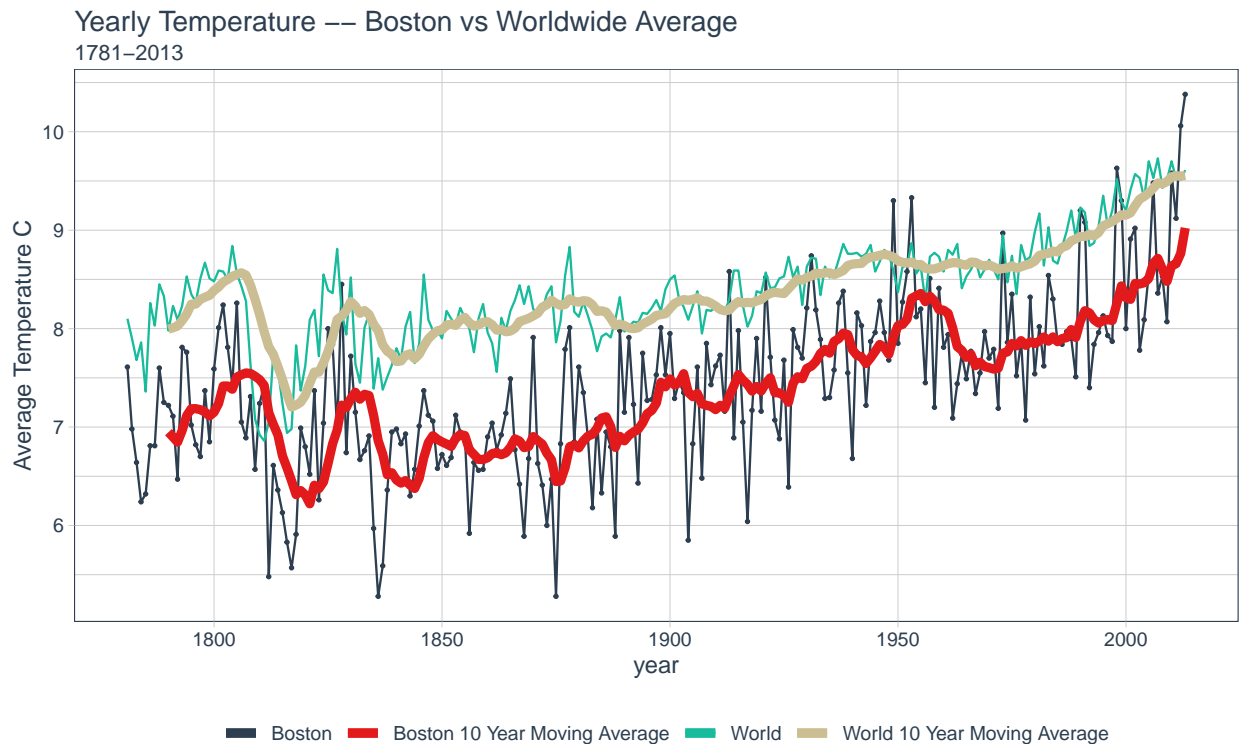Yearly Temperature –– Boston vs Worldwide Average
Boston data noisy prior to 1780, with 1780 missing.
 Remove data prior to 1780 from future analysis
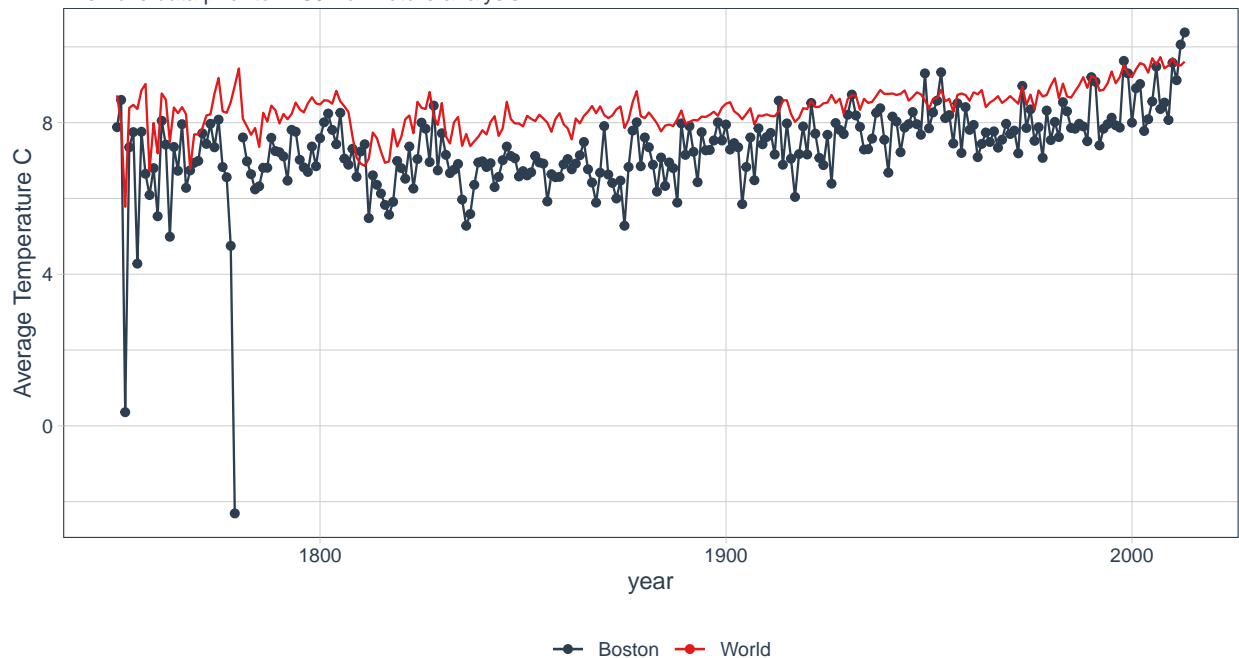
## Data manipulation - Create 10 Year moving average for Boston and World temperatures.

– The window of 10 years is fairly arbitrary, but I chose it to smooth noise that could possibly be interpreted as seasonality in the Boston series. I used the tq_mutate function from the tidyquant package to calculate the moving average.

– Boston series is noisy prior to 1781, so I will drop observations prior, since we have a nice long series to observe.

View Boston and World temperature time-series with 10-year moving averages

Yearly Temperature –– Boston vs Worldwide Average
1781–2013

Legend: Boston — Boston 10 Year Moving Average — World — World 10 Year Moving Average

– We can see in the chart above that the average temperature in Boston is correlated over time with the average worldwide temperature, as both trend up over time. The relationship is smoother for the 10 year moving averages, as yearly fluctuations are smoothed.

– We can see that the Boston temperature series has a larger standard deviation (0.86) than the worldwide temperature (0.54).

– While Boston's yearly temperature tends to be lower than the worldwide average temperature, this is not always the case. Boston had a higher average temperature than the worldwide average for both 2012 and 2013, as well as a handful of other years. As far as this data set goes, the 10 year moving average provides the stability to always keep Boston's 10 year moving average well below that of the world's.

Plot of pairwise comparison and distribution of fields.
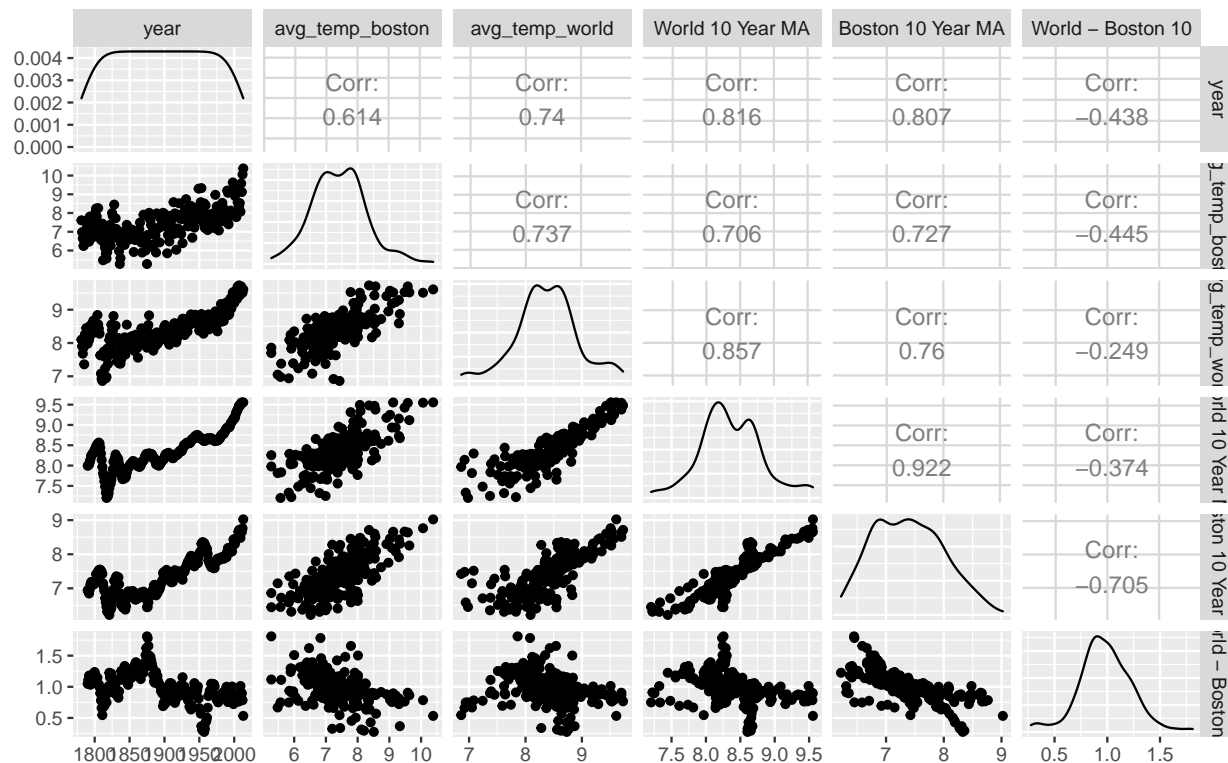
– The box in the upper-right corner of the mosaic above tells us that the difference between the 10 year moving average of the world and Boston is decreasing over time, with a correlation of -0.438 with time. The bottom right square displays a scatter plot of the relationship, suggesting that it is non-linear over time as the World-Boston gap increased in the 19th century before trending downward in the 20th century and beyond.